IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:) Group Art Unit: 2645
)
KRACK, Mike   ) Examiner: Gauthier, Gerald
)
Serial No.: 09/917,576   ) DECLARATION OF MIKE KRACK
) UNDER 37 C.F.R. §1.131
Filed: July 27, 2001   )

Atty. File No.: 4366-37   )
)
For: "METHOD OF PROVIDING   )
      SPEECH RECOGNITION FOR IVR   )
      AND VOICE MAIL SYSTEMS"   )

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Dear Sir:

I, Mike Krack, declare as follows:

1.   I am the inventor of the above-referenced patent application and am familiar with the application. This Declaration is being submitted in connection with prosecution activities for the above-referenced patent application.

2.   From March 1991 until its acquisition by Octel Communications ("Octel"), I was employed by Compass Technologies, Inc. Octel was later acquired by Lucent Technologies ("Lucent"). Lucent later spun off Avaya, Inc. Since its spinoff, I have been employed by Avaya, Inc., as a software engineer. I am also a shareholder of Avaya, Inc.

3.   Attached as Exhibits "A" and "B" are documents generated by myself before the January 10, 2001, filing date of U.S. Patent Application Publication 2002/0090066 to Gupta et al. The document attached as Exhibit "A" is entitled "Patent Submission" and the document attached as Exhibit "B" is a drawing entitled "SEGA: Speech Enabling Gateway System

Overview". Exhibit "A" was prepared on January 3, 2001, and Exhibit "B" predates Exhibit "A".

4.   With reference to independent claims 1, 7, 17, and 29, Exhibit "A" describes an interactive voice response system for a telecommunications system. The system includes an adjunct processor that outputs an output data stream to a user (legacy IVR and Voice Mail Systems) and a speech gateway enabling system (SEGA). As set forth in Exhibit "A", SEGA includes a speech recognition engine operable to identify words in an input voice stream received from the user on a first communication path extending between the user and SEGA. SEGA calls the adjunct processor on a second communication path and conferences the two lines together. SEGA listens for spoken commands on the caller's telephone line (first communication path) and plays the appropriate DTMF digits on the second telephone line (second communication path) connected to the adjunct processor. SEGA maps the spoken commands to DTMF digits. SEGA then plays the DTMF tones corresponding to the spoken commands to the adjunct processor. Because the two lines are conferenced together, the input voice stream, received from the user, is transferred from the first communication path to the second communication path extending between SEGA and the adjunct processor. In this manner, DTMF tones generated by the caller will be played to the adjunct processor. SEGA also transfers the input voice stream from the first communication path to the speech recognition engine for word recognition.

5.   With reference to independent claims 1, 7, 17, and 29, Exhibit "B" discloses a PBX, SEGA (including a Brooktrout™ speech recognition card), and VFS (or adjunct processor). A caller first dials the SEGA hunt group, second SEGA dials the VFS hunt group, and third SEGA connects to the VFS. Stated another way, channel 1 connects the caller to SEGA while channel 2 connects SEGA and VFS. SEGA starts recording on channel 2 and plays

everything recorded from channel 2 on channel 1 in 100 msec blocks. SEGA listens for speech input on channel 1 and, when received, passes the speech from channel 1 to a speech recognition engine resident on the same PC. Output from the speech recognition engine is mapped to one or more DTMF digits. Different "Speech-to-DTMF" tables exist for each VFS. When "Record" or "New Message" is spoken, SEGA begins to record on channel 1 and plays it back on channel 2. Speech recognition is disabled until a DTMF digit is detected on channel 1. The playback from channel 2 is echo canceled, permitting "barge-in".

6. A few weeks before Exhibit "B" was prepared, I contacted Alan Percy of Brooktrout, Inc., for assistance in building a prototype of the system shown in Exhibit "B". I required his assistance in answering my questions about the speech recognition engine sold under the tradename Nuance™. At the time of the contact, a nondisclosure agreement was in place between Brooktrout, Inc., and Lucent Technologies, Inc., my employer at that time. Within two and one half months of the creation of Exhibit "B" and before the filing date of Gupta et al., I had source code written and the prototype up and running, and it was being used and tested internally by the Milpitas, California, engineering group. The prototype falls within the scopes of independent claims 1, 7, 17, and 29.

7. The Patent Submission attached as Exhibit "A" was converted and filed as the above application within seven months of the date of its preparation. It was received by the Legal Department of Lucent Technologies, Inc., on January 5, 2001. After consideration and approval by the patent committee, a case record listing was generated for the Patent Submission on March 6, 2001. The Patent Submission was forwarded to outside patent counsel on March 7, 2001. Outside patent counsel received the Patent Submission on March 8, 2001, and interviewed me about the invention in late May or early June, 2001. A draft application was forwarded by

outside counsel to me on July 6, 2001. After I reviewed the patent application and provided my comments to the responsible patent attorney, I received a revised draft application incorporating my comments on July 20, 2001. The patent application was filed on July 27, 2001.

8.   The foregoing statements and attached exhibits establish conception and actual reduction to practice dates before the January 10, 2001, filing date of Gupta et al. and diligence between the date of Exhibit "B" and the actual reduction-to-practice date and between the date of Exhibit "A" and the constructive-reduction-to-practice date.

7.   I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that the statements were made with the knowledge that willful false statements and the like, so made, are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the subject application or any patent issuing thereon.

Date: _Feb. 21, 2005_____   By:_____
                                                          Mike Krack

J:\4366\-37\Declaration - Mike Krack.doc

-4-

Lucent Technologies
Bell Labs Innovations
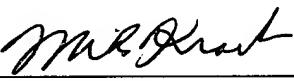
subject: **Patent Submission -** *(Method of providing speech recognition for IVR & Voice Mail systems)*	date: **Jan 3, 2001**

from: **Mike Krack**
Avaya, Inc.
Milpitas, CA
408-324-4378

EMAIL TO: Barry Freedman (Eastern US and Non-US employees)
freedman@lucent.com

or

Dave Volejnicek (US employees in Illinois and west)
czech@lucent.com

COPY TO: Stuart Waldman (all employees)
swaldman@lucent.com

>>>>>>>>>>>>PLEASE LIMIT YOUR SUBMISSION TO 2 to 3 PAGES.<<<<<<<<<<<<
IF YOU HAVE ANY QUESTIONS OR WOULD LIKE A SAMPLE OF A COMPLETED SUBMISSION
FORM, PLEASE CONTACT:

Stuart Waldman
(732) 817-4153
HO 1B-420
swaldman@lucent.com

SUBJECT OF SUBMISSION:

Method of providing speech recognition for DTMF-based IVR & Voice Mail systems.

OBJECTIVE:

Allow speech recognition technology to control legacy IVR & Voice Mail systems which accept only DTMF input.

BRIEF DESCRIPTION:

This solution uses an intermediate Speech Enabling Gateway (SEGA) to provide a speech recognition front-end to any existing DTMF-based Telephone User Interface (TUI). Callers dial into the SEGA system instead of the legacy IVR/VM system. SEGA calls the IVR/VM system on a second telephone line and conferences the two lines together. SEGA

*listens for spoken commands on the caller's telephone line and plays the appropriate DTMF digits on the second telephone line connected to the IVR/VM system. The caller does not hear the DTMF digits as the caller is placed on hold while the DTMF tones are played, creating the a speech enabled TUI. The grammar used by the SEGA system defines the mapping between spoken command and DTMF digit(s). For example, on some voice mail systems, the command "delete" may be mapped to DTMF digits "*3". The caller says "delete message" but the voice mail system hears "*3".*

COMPARISON:

Speech enabled TUIs typically interface with speech recognition software to obtain data describing the spoken commands. The software then branches to the appropriate code depending on the spoken command. The SEGA solution does not require any software changes to existing legacy DTMF-based TUIs. The TUI continues to respond only to DTMF input. The translation from spoken word to DTMF is performed inside the SEGA system and the legacy system is unaware of how the DTMF was generated. As such, it works with any DTMF-based TUI.

This solution provides for a DTMF fallback in the event that the speech software is unable to accurately recognize the spoken commands. Since the two channels are conferenced, any DTMF generated by the caller will be played to the IVR/VM system on the second telephone line as well.

USE:

Several companies offer speech recognition TUI's, but I haven't heard of any products to speech enable legacy DTMF-based TUIs.

As demand for 'hands-free' telephone operation grows, speech recognition could very likely become an expected feature of voice mail access products .

This feature would also work well on DTMF-based IVR systems, such as Avaya's Conversant applications.

This approach can use VoiceOverIP technology to provide a software only solution, where H.323 replaces the functionality provided by the 2 analog telephone lines described above.

A working prototype has been up and running and used internally by the Milpitas,CA engineering group since '   It was functional in a lab environment 1 week earlier and first described on an internal Avaya web site on   See http://148.147.189.20/sega.htm or http://135.74.136.91/sega.htm for more info.

ECONOMIC IMPACT:

As demand for 'hands-free' telephone operation grows, speech recognition could very likely become an expected feature of voice mail access products. Avaya Inc alone has 100 million voice mailboxes that could be speech enabled by this solution.

This feature would also work well on DTMF-based IVR systems, such as Avaya's Conversant applications.

FOREIGN INTEREST:
(1. In which foreign countries, if any, should we obtain a patent? Why (e.g., big market there; major competitors are based there)?)
FILL IN

PUBLICATION OF PROPOSED INVENTION: (List publications, if any, which have described the proposed invention. Include dates.)
FILL IN

ORIGINATORS OF THE PROPOSAL:
(Name, Dept., Room number, Ext.)

Mike Krack, mkrack@avaya.com 408-324-4378

CLASSIFICATION:
(Based on the Lucent Patent Filing and Maintenance Policy chart (attached), what is the Classification Code for this idea:
Your rating? ___II – *Important (could become a standard feature and method)* _____

Your department Head's Name and Rating? _Bill Mccarty_____

CONCEPTIONS OF INVENTION:

Date of first drawing(s):   _   _____

Where is drawing located?: _http://148.147.189.20/sega.htm__

Date of first written description: _.   _____

Where is description located?: __ http://148.147.189.20/sega.htm

Date of first oral discussion to others: __   _____

To whom?: _Doug Zumbiel (dougz@avaya.com), Russ Innes (rinnes@avaya.com), Ray Barbieri (rayb@avaya.com)
Attachments:

SEGA: *Speech Enabling Gateway*
System Overview

PBX
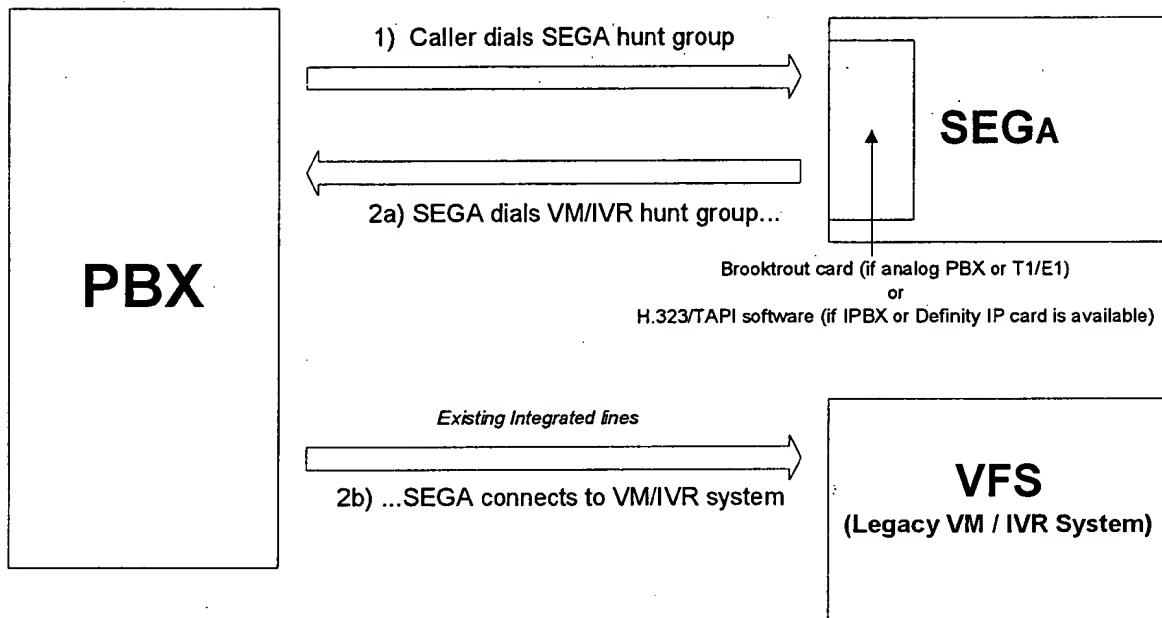
1) Caller dials SEGA hunt group →

← 2a) SEGA dials VM/IVR hunt group...

SEGA

Brooktrout card (if analog PBX or T1/E1)
or
H.323/TAPI software (if IPBX or Definity IP card is available)

Existing Integrated lines →

2b) ...SEGA connects to VM/IVR system

VFS
(Legacy VM / IVR System)

1) Caller dials a new hunt group number, which goes to SEGA (ch. 1).
2) SEGA answers on ch. 1 and calls the old VM/IVR hunt group on ch. 2.
3) SEGA starts recording on ch. 2.
4) SEGA plays everything recorded from ch.2 on ch. 1, in 100 msec blocks.
5) Any DTMF detected on ch 1 is generated on ch. 2.
6) SEGA listens for speech input on ch. 1 (the playback from ch.2 is echo canceled, permitting 'barge-in') *
7) SEGA passes speech input from ch. 1 to a speech recognition engine resident on the same PC. Output from the engine is mapped to a DTMF digit(s). Different "Speech-to-DTMF" tables exist for each VFS. i.e. "Delete = *D" or "Delete = 337"

* When "Record" or "New Message" is spoken, SEGA begins to record on ch. 1 and plays it back on ch. 2. Speech recognition is disabled until a DTMF digit is detected on ch. 1.